TITLE OF THE INVENTION:

FLOW CONTROL AND CONGESTION CONTROL IN A DATA DISTRIBUTION NETWORK

CROSS REFERENCE TO RELATED APPLICATIONS:

[0001] This specification claims priority from U.S. Provisional Patent Application Serial No. 60/483,025, entitled "Flow Control and Congestion Control in XGS," and filed on June 27, 2003, and U.S. Provisional Patent Application Serial No. 60/529,618, entitled "Flow Control and Congestion Control in a Data Distribution Network," and filed on December 16, 2003. The contents of the above-referenced Provisional Patent Applications are hereby incorporated in their entirety by reference.

BACKGROUND OF THE INVENTION:

Field of the Invention:

[0002] Certain embodiments of the present invention are directed generally to methods for managing flow of datagram traffic across a data distribution network. Certain other embodiments of the present invention are directed generally at methods for controlling congestion in data distribution networks. Further, additional embodiments of the present invention are directed generally to devices and systems for implementing the above-discussed methods.

Description of the Related Art:

[0003] According to modern telecommunications networks, datagrams such as, for example, packets, cells, and bit strings, can be commonly distributed from one node, module, or device (generally referred to herebelow as network devices) on a network to one or more other network devices. During the distribution of such datagrams, the ingresses and egresses of the network devices that assist in the distribution and/or forwarding of

datagrams often get congested or clogged due to over-subscription. Two examples of how over-subscription can affect a data distribution network device are shown in Figures 1 and 2, respectively.

[0004] Figure 1 illustrates device D1 and device D2 that are operably connected to each other. Device D1 has ports P1-P6 that may each function as ingresses and/or egresses of datagrams from and/or to other network device (not illustrated in Figure 1) into device D1. Device D1 also has a port P7 through which datagrams in device D1 may egress device D1 towards device D2 and through which datagrams in device D2 may ingress into device D1.

[0005] Port P7 on device D1 is operably connected to port P8 that may function as an ingress for device D2 illustrated in Figure 1. Device D2 includes additional ports P9-P14 through which datagrams that enter through port P8 may egress from device D2 and through which device D2 may communicate with other network devices (not illustrated in Figure 1).

[0006] According to the related art devices illustrated in Figure 1, port P5 on device D1 and port P14 on device D2 may both wish to simultaneously transmit datagrams to port P13 on device D2. Under such circumstances, datagrams from port P5 travel through port P7 of device D1 and through port P8 of device D2 before arriving at port P13.

[0007] When, as shown in Figure 1, a pathway 100 between port P7 and port P8 is capable of supporting a flow of datagrams higher than either port P5 or port P13 would be able to accommodate individually, pathway 100 between port P7 and port P8 may be used as or replaced with a high-speed and/or high-capacity pathway that can support a high volume of datagram traffic. In fact, under ideal situations, a high-speed/capacity pathway is able to support flow of datagrams between many different ports on device D1 and device D2, all simultaneously.

[0008] It should be noted that the pathways discussed herein are generally not physical connections. Rather, these pathways often are made up of a series of data distribution devices, such as, for example, routers and switches, that, at least for a brief period of time, allow a datagram to travel from a source to a destination on a network.

[0009] Because, in Figure 1, ports P1-P6 and P9-P12 all have the same capacity for allowing datagrams to ingress/egress devices D1 and D2, port P13 is over-subscribed because it is simultaneously receiving full-streams of datagrams from both port P14 and port P5, thereby exceeding port P13's capacity for egressing received datagrams. Once this over-subscription condition on P13 is detected in device D2, according to the related art, a "pause" frame is sent to port P8 and port P14 within device D2.

[0010] Once port P8 receives the "pause" frame, the pathway between related art network devices D1 and D2 is effectively blocked, at least for a short time interval. This is true because, in the "paused" condition, port P8 will not accept datagrams from port P7.

[0011] Once the pathway is effectively blocked between devices D1 and D2, if port P2 on device D1 and port P10 on device D2 try to exchange datagrams, the datagram exchange would not be possible. Hence, even though, in the related art example illustrated in Figure 1, neither port P2 nor port P10 are over-subscribed, and even though there may be available bandwidth on the pathway between port P7 and port P8 which link devices D1 and D2, the "pause" frame emitted as a result of the over-subscription of port P13 has affectively shut down all communications between devices D1 and D2.

[0012] Hence, there is a need for devices, systems, and methods that prevent the over-subscription of one port on one network device from effectively shutting down either a portion of or an entire data distribution network. In other words, there is a need for better flow control of datagrams over data distribution networks.

[0013] Figure 2 illustrates how related art devices D1 and D2 illustrated in Figure 1 deal with an over-subscription condition on one their ports in a different manner. As in Figure 1, port P13 in Figure 2 is again over-subscribed because port P13 is receiving more datagrams from port P14 and port P5, via port P7 and port P8 that are connected via a high-speed/capacity pathway, than P13 can allow to egress.

[0014] In contrast to Figure 1, in the over-subscribed condition illustrated in Figure 2, port P13 does not send out a "pause" control frame. Rather, the first device D1 detects the over-subscription of port P13 and reacts accordingly by only allowing portions of the datagrams sent from ports P14 and P5 to reach port P13.

[0015] More specifically, in a situation illustrated in Figure 2 where, for example, port P5 on device D1 is transmitting one Gigabit ("one Gig") of datagrams, port P14 on device D2 is transmitting one Gig of datagrams, and port P13 is only capable of egressing one Gig of datagrams, port P8 allows only, for example, half of the datagrams received from port P5 to travel to port P13. In such an example, if port P14 is similarly restricted (generally by internal components of the device D2) to only being able to forward half a Gig of datagrams to port P13, port P13 will receive the one Gig of datagrams that it can egress, and half of the datagrams received at port P8 will be dropped/lost at port P8.

[0016] Clearly, dropping of datagrams anywhere in a data distribution network is a non-ideal and non-desirable condition. Further, it is even more non-desirable for an extra half Gig of datagrams to travel between port P7 and port P8 when this half Gig of datagrams cannot be supported by port P13. This is true at least because the extra datagrams only serve to additionally congest the pathway between device D1 and device D2, which could be more effectively used by datagrams having other sources and destinations.

[0017] Hence, what is also needed are methods and devices for reducing the amount of congestion that is present in high-speed/capacity pathways between network devices. In addition, there is also a need for methods and devices that can halt the forwarding of unusable datagrams as early as possible.

SUMMARY OF THE INVENTION:

[0018] According to certain embodiments of the present invention, a method of operating a telecommunications network is provided. Typically, this method includes the step of providing a first networked device that is operably connected to a second networked device. This method also commonly includes transferring datagrams from a first port of the first device to a first port of the second device using a pathway that is operably connected to a second port of the first device and a second port of the second device. In addition, this method usually includes the step of selectively pausing an individual port on the first device that is causing over-subscription of the first port of the second invention.

[0019] According to certain other embodiments of the present invention, another method of operating a telecommunications network is provided. A common first step of this method includes providing a first networked device that is operably connected to a second networked device. Typically, this method also typically includes the step of transferring datagrams from a first port of the first device to a first port of the second device using a pathway that is operably connected to a second port of the first device and a second port of the second device. This method also generally includes signaling the first port of the first device to send fewer datagrams to the first port of the second device when an over-subscription is detected at the first port of the second device.

[0020] Certain embodiments of the present invention also provide a communication system. According to some of these embodiments, a first

data distribution means that is operably connected to a second data distribution means is included. Often, a first communication means, for transferring datagrams from a first port of the first data distribution means to a first port of the second data distribution means, is also provided. Further, a control means for selectively pausing individual ports that are causing over-subscription of the first port of the second data distribution means is generally provided.

[0021] According to yet other embodiments of the present invention, another communications system is provided. Typically, this system includes a first data distribution means that is operably connected to a second data distribution means. Both of these means normally may be used for distributing datagrams over a network. Generally, this system also includes communications means for transferring the datagrams from a first port of the first data distribution means to a first port of the second data distribution means. In addition, this network usually includes a control means for signaling the first port of the first data distribution means to send fewer datagrams to the first port of the second data distribution means when an over-subscription is detected at the first port of the second data distribution means.

[0022] According to other embodiments, yet another communications system is provided. In this system, a first device that is operably connected to a second device is usually provided. Also, a first controller capable of transferring datagrams from a first port of the first device to a first port of the second device are typically included. Further, a second controller capable of selectively pausing individual ports in the first device that are contributing to over-subscription of the first port of the second device are normally included.

[0023] In addition to the above, certain other embodiments of the present invention provide still another communications system. According to this communications network, a first device that is operably connected to a

second device is typically provided. Also, a first controller capable of transferring datagrams from a first port of the first device to a first port of the second device is often provided. In addition, a second controller capable of signaling the first port of the first device to send fewer datagrams to the second port of the second device when an over-subscription is detected at the second port of the second device is usually provided.

BRIEF DESCRIPTION OF THE DRAWINGS:

[0024] For a proper understanding of the invention, reference should be made to the accompanying drawings, wherein:

[0025] Figure 1 illustrates a portion of a related art data distribution network wherein an over-subscription of one port in the network may cause pausing of the entire network;

[0026] Figure 2 illustrates the portion of the related art data distribution network illustrated in Figure 1 wherein over-subscription of one port in the network may lead to unnecessary congestion of a high-capacity network pathway;

[0027] Figure 3A illustrates two network devices according to certain embodiments of the present invention wherein a port on one of the devices is over-subscribed and Figure 3B illustrates an embodiment of the present invention wherein the over-subscription condition illustrated in Figure 3A has been addressed by selectively signaling an offending port that the offending port should alter the number of datagrams forwarded to the over-subscribed port;

[0028] Figures 4A and 4B illustrate certain embodiments of the present invention wherein an over-subscription of one port in a network is addressed by signaling all other ports in at least a portion of the network that all ports should to alter the number of datagrams sent to the over-subscribed port;

[0029] Figure 5A illustrates a flow chart that includes steps of methods according to certain embodiments of the present invention for distributing datagrams under a flow control mechanism;

[0030] Figure 5B illustrates a flow chart that includes the steps of a method according to certain embodiments of the present invention wherein a network that has been paused while operating under a flow control mechanism is resumed;

[0031] Figure 6A illustrates a flow chart that includes the steps of a method according to the certain embodiments of the present invention wherein a congestion control mechanism is used;

[0032] Figure 6B illustrates a flow chart that includes the steps of a method according to certain embodiments of the present invention wherein a network that has been head of line (HOL) blocked while operating under a congestion control mechanism resumes operations;

[0033] Figure 7 illustrates a flow chart that includes the steps of a method according to certain embodiments of the present invention wherein a flow control mechanism pauses and subsequently resumes operation of a network; and

[0034] Figure 8 illustrates a flow chart that includes the steps of a method according to certain embodiments of the present invention wherein a congestion control mechanism blocks and subsequently resumes network operation.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS:

[0035] In order to address and overcome the shortcomings of the related art methods, systems, and devices discussed above, embodiments of the present invention include a variety of methods, devices, and systems have been developed. Some of these are discussed herebelow and illustrated in the attached figures.

[0036] Figure 3A illustrates data distribution device 305 that is operably connected to data distribution device 310. According to certain embodiments of the present invention, each of devices 305 and 310 may be used as a means for distributing datagrams over a network.

[0037] Each device 305, 310 illustrated in Figure 3A includes a set of ports P15-P32, an Interconnect Port Controller (IPC) 315, 320, a memory unit controller 325, 330, and a memory unit 335, 340. Each memory unit 335, 340 illustrated includes a bitmap 345, 350 or other set of instructions.

[0038] Pathway 355 is illustrated between devices 305 and 310. According to certain embodiments of the present invention, pathway 355 acts as a communication means that allow for communication between devices 305 and 310. Pathway 355 may provide for non-lossy communication between devices 305 and 310, wherein the loss/dropping of datagrams that travel between devices 305 and 310 is eliminated, or at least greatly reduced. According to certain embodiments, pathway 355 may be a physical or logical connection.

[0039] As shown in Figure 3A, when in operation, port P20 of device 305 may forward datagrams to memory unit controller 325. Controller 325 generally references bitmap 345 in memory unit 335 operably connected thereto before forwarding the datagrams to IPC 315 of data distribution device 305. Then, using pathway 355 that effectively connects ports P23 and P24, the datagrams may be forwarded to IPC 320 of data distribution device 310.

[0040] The datagrams that reach IPC 320 are then usually forwarded to memory unit controller 330, which commonly references bitmap 350 in memory unit 340 of data distribution device 310. Then, these datagrams may be forwarded to port P31. Also, when the portion of the network illustrated in Figure 3A is in operation, port P32 may forward datagrams to port P31.

[0041] Figure 3B illustrates how devices 305 and 310 of Figure 3A respond, according to certain embodiments of the present invention, when the combination of the datagrams received at port P31 from port P20 and port P32 results in the over-subscription of port P31. It should be noted that, in Figure 3B, memory unit controller 330 typically stores and/or maintains threshold values for every ingress port in devices 310 and 305. Hence, when port P20, which is an ingress port while it is sending datagrams to port P31, exceeds a threshold value maintained by memory unit controller 330, an over-subscription condition is detected.

[0042] Because the offending port in the situation above is ingress port P20, port P31 will typically be able to continue to transmit datagram traffic ingressed at ports other than port P20. However, when the above-discussed over-subscription is detected, memory unit controller 330 typically sends one or more pause frames to port P20, the ingress port that exceeded the threshold, thereby at least temporarily preventing port P20 from forwarding any additional datagrams.

[0043] In other words, according to certain embodiments of the present invention, port P20 on device 305 that contributes to the over-subscription of port P31 on device 310 may be paused selectively, without having to pause any other ports on the network. According to such embodiments, if port P16 is instructed to forward datagrams to port P27, such transmission would be possible, so long as there is adequate bandwidth available along pathway 355, even when one of the ports on the network is over-subscribed.

[0044] It should be noted that all of the controlling of ports discussed herein may be effectuated using non-memory-consuming methods of communication. Using such non-memory-consuming methods allows for information related to how individual ports should operate to flow across the network independently of network congestion.

[0045] A non-memory-consuming transaction generally does not make use of memory resources, such as the resources illustrated in memory units 335

and 340. For example, a flow control frame may be sent directly from IPC 320 to IPC 315 over pathway 335. In a non-memory-consuming transaction IPC 315 may then receive the frame and decode the frame's contents, construct messages, and update a port's state in memory unit 335 which, in turn, may take any further action necessary, such as informing port P20 to send a pause frame.

[0046] In the above-described situation, if the control frame were memory-consuming, the reception of the frame would be dependent on the memory resource available in pathway 335. However, using a non-memory-consuming transaction such as the transaction described above generally guarantees that information contained in the frame will be conveyed, regardless of the memory resources available.

[0047] Figure 4A again illustrates above-discussed data distribution devices 305 and 310. In Figure 4A, the forwarding of datagrams by port P38 and port P50 to port P49 results in the over-subscription of port P49. However, as will be shown in Figure 4B, a congestion control mechanism is used to address the over-subscription issue.

[0048] According to Figure 4B, which illustrates a possible response mechanism for an over-subscription condition according to certain embodiments of the present invention, over-subscribed port P49 informs memory unit controller 330 of the over-subscription. Then, after updating bitmap 350 in memory unit 340 to which it is operably connected, memory unit controller 330 broadcasts to all ports within data distribution device 310 that port P49 is over-subscribed and instructs all of these other ports to either send no more datagrams to port P49 or to reduce the amount of datagrams sent to port P49 in order to eliminate the over-subscription.

[0049] According to certain embodiments, memory unit controller 330 also informs data distribution device 305, via IPC 320 and pathway 355, of the over-subscription. Data distribution device 305 then uses IPC 315 and memory unit controller 325 to update bitmap 345 to reflect the over-

subscription of port P49 and to communicate with the ports on data distribution device 305 that they should either stop forwarding datagrams to port P49 or decrease the number of datagrams sent to port P49 in order to eliminate the over-subscription thereof.

[0050] In Figure 3A, 3B, 4A, and 4B, controllers 320, 315, 325 and 330 and pathway 355 may, according to certain embodiments of the present invention, form communications means for transferring datagrams from port P20 on data distribution device 305 to port P31 on data distribution device 310. According to certain embodiments of the present invention, all of controllers 315, 320, 325 and 330, or any individual controller or subset of these controllers, may be replaced by, for example, a more general controller capable of transferring datagrams from data distribution device 305 to data distribution device 310.

[0051] Memory means 335 and 340 and/or bitmaps 345 and 350 may, according to certain embodiments, be used as storage means for storing information concerning which ports in the network are over-subscribed. Also, although bitmaps 345 and 350 are illustrated in Figures 3A, 3B, 4A, and 4B, these, according to certain embodiments, may be replaced by other formats for storing information related to whether ports on the network are over-subscribed and/or for storing information concerning how much information may be sent to various ports on the network without causing over-subscriptions.

[0052] According to certain embodiments, controllers 315, 320, 325 and 330 may be replaced with more general control means for selectively pausing or curtailing the amount of datagrams sent by individual ports in the network that are causing over-subscriptions. According to certain embodiments, these controlling means may also be in the form of actual controller devices.

[0053] Figure 5A illustrates flowchart 500, which illustrates the steps of a method, according to certain embodiments of the present invention, wherein

a network such as, for example, a telecommunications network, is operated. In order to illustrate a representative implementation of the method whose steps are illustrated flowchart 500, reference will be made to devices 305 and 310, which make up the portion of the network illustrated in Figure 3A.

[0054] According to step 505, illustrated in Figure 5A, a second device receives a datagram from a first device through an interconnect port. In Figure 3A, step 505 may, for example, be implemented when a datagram originating from port P32 in first data distribution device 305 arrives at port P24 in second data distribution device 310.

[0055] Step 510 recites maintaining a first threshold for one or more ingress ports other than the interconnect port. This first threshold may be maintained, for example, in a memory control unit that is operably connected to the interconnect port and is typically a high threshold that specifies a maximum number of datagrams that may be supported. According to certain embodiments of the present invention, memory control units in each device on a network may maintain individual high thresholds, or high threshold values, for all ingress ports beyond the interconnect ports of their respective devices. In Figure 3A, memory units 335 and 340 may each individually maintain and/or store individual high threshold values.

[0056] Step 515 recites accounting for each datagram that originated over the interconnect port by incrementing an ingress datagram count. Step 515 may, for example, be implemented by providing a counter in a first network device and incrementing the value of the counter each time that a port that provides a connection with other network devices receives a datagram. In Figure 3A, a counter may be stored, for example, in IPC 320 of second data distribution device 310. Then, every time that port P24 receives a datagram via pathway 355, the value of the counter may be incremented.

[0057] The value of the counter is typically decremented when a datagram egresses from the device that had received the datagram. Hence, in Figure 3A, once a datagram ingresses into second data distribution device 310

through port P24 and the counter is incremented, the counter may be decremented once the datagram egresses though any of ports P25-P32.

[0058] Step 520 recites comparing the ingress datagram count against the first or high threshold value to determine whether an over-subscription has occurred at the ingress port. Also, step 525 recites directing a pause control frame specifically at the port that caused the over-subscription when an over-subscription has occurred. In Figure 3, step 520 may be implemented by repeatedly comparing a high threshold value that specifies whether port P24 is over-subscribed with an ingress datagram count that is incremented each time that a datagram ingresses into second data distribution device 310 through port P24. If the comparison shows that the count is greater than or equal to the high threshold value, an over-subscription is detected. Then, upon detection of this over-subscription, a pause control frame may be sent to whichever network port sent the datagram that caused the count to reach or exceed the high threshold.

[0059] The above-described set of steps, performed either individually or in combination, may be used to prevent over-subscription in network devices by pausing specific ports on the network. However, once a port has been paused, it is typically desirable to allow the paused port to resume transmitting datagrams once the over-subscription condition has been eliminated. Hence, Figure 5B illustrates flowchart 530, which includes the steps of a representative method wherein the paused port discussed above in connection with Figure 5A is allowed to resume forwarding datagrams.

[0060] Step 540 in Figure 5B recites maintaining a second threshold for one or more ingress ports other than the interconnect port. Like the first threshold, this second threshold may be maintained, for example, in a memory control unit that is operably connected to the interconnect port. Unlike the first threshold, however, this second threshold is typically a low threshold limit, at which point the memory control unit determines that there

are enough resources available for the ingress port to receive a datagram, as recited in step 550.

[0061] Once this second or low threshold is reached, as recited in step 560, the memory control unit may construct a "resume" frame and send it, specifically, to that ingress port that had previously been paused.

[0062] According to certain embodiments of the present invention, the re-activating of a paused port may be performed by transmitting a re-activation signal in the form of, for example, a datagram or control frame, to the paused port. According to certain embodiments, the re-activating of a paused port may be performed pursuant to the detection of a condition wherein the first port of the second device has datagram traffic flowing therethrough in an amount that is below a lower trigger value. For example, if a first port capable of supporting 1 Gig of datagram traffic becomes over-subscribed and leads to the pausing of a second port, datagram transmission from the second port may, according to certain embodiments, resume after the over-subscription of the first port has ended and only a half Gig of datagrams are being forwarded to the first port.

[0063] According to certain other embodiments, the re-activating of a paused port may occur pursuant to the passage of a pre-determined amount of time. For example, if a first port becomes over-subscribed and causes the pausing of a second port, the second port may be re-activate after a half-second time period has passed.

[0064] Figure 6A illustrates flowchart 600 that includes the steps of a representative method according to certain embodiments of the present invention wherein a congestion control mechanism is used. According to the step 605 in Figure 6A, a second device receives a datagram. According to certain embodiments, the datagram is received from device 305 illustrated in Figure 3A, but any other portion of the network that is operably connected to the device 310 illustrated in Figure 3A may also send the datagram.

[0065] Step 610 in Figure 6A specifies that a memory unit of the second device should determine that an egress port of the second device to which the datagram has been sent cannot accommodate the received datagram. Then, step 615 in Figure 6A specifies that the memory unit of the second device should update a head of line (HOL) blocking bitmap to include therein that the datagram cannot be accommodated by the egress port. In Figure 3A, bitmap 345 and/or bitmap 350 may be used as an HOL bitmap. It should also be noted that, according to certain embodiments, in addition to the bitmap that is specifically recited in step 615, any other format that stores information related to which ports are blocked/over-subscribed may be used.

[0066] Step 620 of Figure 6A specifies that a memory unit of the second device should inform the interconnect port controller (IPC) of the second device of the updated bitmap. According to certain embodiments, after updating an HOL bitmap, a memory unit notifies the IPC to which it is operably connected of the updated HOL bitmap.

[0067] According to step 625 in Figure 6A, the IPC of the second device should construct a broadcast control frame and should thereby inform other network devices that there is HOL blocking for the egress port of the second device discussed above. According to certain embodiments, the IPC sends out a control frame to all devices and/or ports on the network and informs them all that there is congestion/over-subscription at a port specified within the control frame.

[0068] According to step 630 of Figure 6A, the IPC on other network devices should interpret the control frame and should update the bitmaps on the other devices to which the control frame was sent. Generally, this step prevents more datagrams from being sent to the egress ports of the second device specified above as being congested/over-subscribed.

[0069] Figure 6B illustrates a flowchart that includes the steps of a representative method for ceasing the HOL blocking that was used according to the method illustrated in Figure 6A. In Figure 6B, step 640 recites that the

memory unit of the second device should determine that the egress port of the second device can resume receiving datagrams. Step 640 also specifies that the memory unit should update the HOL bitmap and inform the IPC of the second device of the determination made by the memory. It should be noted that, although a memory unit is specifically recited in Figure 6B as making the detection, the detection may be made at a variety of other locations, either within or exterior to the network.

[0070] Step 645 in Figure 6B specifies that the IPC of the second device should broadcast a control frame indicating that the HOL blocking discussed in Figure 6A has been ceased/removed. Then, step 650 in Figure 6B specifies that the IPC of other network devices should interpret the control frame and allow for datagrams to be sent to the egress port of the second device, is necessary/desired.

[0071] All of the methods for resuming datagram transmission from the paused ports discussed above may be used when ceasing/removing of HOL blocking of ports as well. For example, the ceasing/removing of HOL blocking of ports may occur pursuant to the sending of an appropriate datagram or control frame, the passage of a pre-determined amount of time, or after a lower trigger value has been reached.

[0072] Figure 7 illustrates yet another method according to the certain embodiments of the present invention wherein portions of a network are paused and then resumed or reactivated. According to step 710, a first network device that is operably connected to a second networked device is provided. According to step 720, datagrams are transferred from a first port of the first device to a first port of the second device using a pathway that is operably connected to a second port of the first device and a second port of the second device. Then, according to step 730, an individual port is selectively paused on the first device, particularly when this port is causing over-subscription of the first port of the second device. Step 740 in Figure 7 then specifies re-activating a paused port.

[0073] Figure 8 illustrates a flowchart including steps of yet another method according to certain embodiments of the present invention wherein HOL blocking may be used and wherein resumption of datagram transmission occurs. According to step 810 illustrated, a first network device that is operably connected to a second network device is provided.

[0074] According to step 820 illustrated in Figure 8, datagrams are transferred from the first port of the first device to a first port of the second device using a pathway that is operably connected to a second port of the first device and a second port of the first device and a second port of the second device. Then, according to step 830, a signal is sent to the first port of the first device to send fewer datagrams to the first port of the second device when an over-subscription is detected at the first port of the second device. Then, according to step 840, unrestricted datagram transmission is resumed to the first port of the second device.

[0075] One having ordinary skill in the art will readily understand that the invention as discussed above may be practiced with steps in a different order, and/or with hardware elements in configurations, which are different than those, which are disclosed. Therefore, although the invention has been described based upon these preferred embodiments, it would be apparent to those of skill in the art that certain modifications, variations, and alternative constructions would be apparent, while remaining within the spirit and scope of the invention. In order to determine the metes and bounds of the invention, therefore, reference should be made to the appended claims.